N91-20663

INVESTIGATION OF VARYING GRAY SCALE LEVELS FOR REMOTE MANIPULATION

John M. Bierschwale, Mark A. Stuart, and Carlos E. Sampaio
Lockheed Engineering and Sciences Company
Houston, TX 77058

ABSTRACT

A study was conducted to investigate the effects of variant monitor gray scale levels and workplace illumination levels on operators' ability to discriminate between different colors on a monochrome monitor. It was determined that 8-gray-scale viewing resulted in significantly worse discrimination performance compared to 16- and 32-gray-scale viewing and that there was only a negligible difference found between 16 and 32 shades of gray. Therefore, it is recommended that monitors used while performing remote manipulation tasks have 16 or above shades of gray since this evaluation has found levels lower than this to be unacceptable for a color discrimination task. There was no significant performance difference found between a high and a low workplace illumination condition. Further analysis was conducted to determine which specific combinations of colors used in this study can be used in conjunction with each other to ensure error-free color coding/brightness discrimination performance while viewing a monochrome monitor. It was found that 92 three-color combinations and 9 four-color combinations could be used with 100% accuracy. The results can help to determine which gray scale levels should be provided on monochrome monitors as well as which colors to use to ensure the maximal performance of remotely-viewed color discrimination/coding tasks.

INTRODUCTION

Telerobotic workstations will play a major role in the assembly of Space Station *Freedom* and later in construction and maintenance in space. For the short-term, control of these telerobotic systems will be dependent primarily on the human operator. Since the human operator will be a part of the telerobotic system, it is critical that the components of this interface be designed so that the human operator's capabilities and limitations are best accommodated within the structure of specific task requirements. To emphasize the importance of a well-designed human-telerobot interface, one study found that the selection of an appropriate control device, based upon the operator's capabilities and the requirements of the task, can more than double the productivity of the telerobotic system (O'Hara, 1986).

One of the most important components of the workstation will be the vision system. During many tasks, where a direct view is not possible, cameras will be the user's only form of visual feedback. Monochrome viewing has been used in the past for remote manipulation and inspection. Even though color operations have been baselined for *Freedom*, special situations may dictate the use of monochrome viewing. Other telerobotics users such as future space systems, space station investigators, the nuclear industry, and undersea industry may find the use of monochrome viewing to be advantageous.

An important system parameter for monochrome viewing is number of gray scales that can be displayed on a monitor. The term "gray scale" refers to the uniform variation from black to white through various shades of gray in a television screen image when cathode ray tube (CRT) control voltages are adjusted over the full range of brightness for a specific monitor. The number of gray scales is an issue because more gray scale divisions require greater band width and processing capabilities which are often limited in remote environments. Therefore, the optimal number of gray scales from an economic or cost effectiveness standpoint will be the smallest number of gray scales that provide acceptable performance by the operator. 8, 16, and 64 gray scale values are being considered by engineers for displays.

A survey of the literature has failed to provide definitive guidelines concerning this issue (e.g., Johnston, 1968; Troy, Deutsch, and Rosenfeld, 1973; Shurtleff, 1980; and Kingdom and Moulden, 1986). Woodson (1980) recommended that a minimum of five gray levels be

used for monitors, but precise quantification of the specific gray scale levels to use was not found. Tannas (1985) stated that contouring in big, tone-changing areas is bad if the number of gray scales is less than 16 and that 64 shades of gray should be used for good, aesthetic picture quality images. Tannas also stated that good gray scale performance requires a pixel contrast ratio of approximately 20:1 with the luminance either continuously variable or controllable into at least 16 logarithmically spaced steps. There is little research on the effect of variant gray scales on the performance of an operator performing a remotely-viewed task.

The computer equipment used in this evaluation had the capability of displaying between 2 and 256 different gray scales. Since it was impossible to compare all levels, the first objective of this study was to determine which of the possible gray scale levels are discriminable from each other.

One telerobotic task which would be performed on *Freedom* would be the assembly or maintenance of a thermal utility connector. Color coding may be used to indicate when the valves on each hose are either fully opened or closed. This would be helpful when the task is viewed on a color monitor or performed by an extravehicular astronaut. If this task were viewed on a monochrome monitor, then the information that is used to discriminate between the two colors is the brightness level of the colors. The second objective of this study was to study the discrimination performance of operators while they viewed colored visual stimuli through a monochrome monitor using the pre-determined number of variant gray scale conditions.

Since lighting conditions vary greatly in outer space, the third objective of this study was to also investigate the effect of different illumination levels on the experimental task. The performance of the color discrimination task took place under two different illumination levels as well as variant gray scales. (For a discussion on the effects of lighting on remote manipulation tasks see Chandlee, Smith, and Wheelwright, 1988).

These three objectives were investigated by conducting two different evaluations. The first evaluation narrowed down the number of gray scales while the second evaluation addressed the remaining two objectives.

EVALUATION 1

The objective of Evaluation 1 consisted of determining the discriminable gray scale levels to be investigated in Evaluation 2.

Subjects

Seven volunteer male subjects were selected to participate in this evaluation. All subjects who participated had their vision tested at the JSC Clinic and it was determined that they all had either corrected or uncorrected 20/20 vision and none had evidence of color deficiencies.

Apparatus

Testing took place in the Remote Operator Interaction Laboratory (ROIL) of the NASA Johnson Space Center. The video signal from a digital camera (focused on an extra-vehicular activity (EVA) toolbox) was processed through a DataCube digital image processing system and displayed on a Conrac monitor. The DataCube's output was a monochrome image with one of the following gray scale levels: 8, 16, 32, 64, 128, and 256. The illumination level in Evaluation 1 was the laboratory ambient lighting level of 269 lx (25 fc).

Procedure

The procedure followed was a paired-comparison psychophysics technique. Each operator was randomly presented the image using one of the six gray scale levels as a reference. The reference was then successively paired to each of the other gray scale levels so that a comparison could be made. The reference gray scale was also paired to itself as a control condition. The pairs were formed serially in this part of the testing. Immediately after each subject viewed each of the reference-comparison pairs, they would then state whether or not there was a perceptible difference between the two stimuli. After completing these six paired-comparisons, another of the remaining five gray scale levels was selected as the next reference. Six more paired-comparisons were then conducted. The procedure continued in this fashion until all six of the gray scale levels had served as the

reference gray scale level for this paired-comparison task.

Results and discussion

The data were analyzed in terms of determining the number of subjects who noticed a difference between a specific gray scale level and the other gray scale levels during the paired-comparisons tasks. If a difference was not noticed, then this was classified as an error.

The discrimination error-rates were then statistically analyzed with an analysis of variance (ANOVA) procedure. Based upon the results of the data analysis, it was determined that there was a significant effect ($p < 0.05$) due to the different gray scale levels used. A Newman-Keuls pairwise comparison statistical procedure was then administered to these data. The Newman-Keuls revealed a consistent trend for subjects to notice a difference between 8 shades of gray and all other levels and between 16 shades of gray and all other levels. This trend was not observed for discriminations with 32, 64, 128, and 256 shades of gray.

Based upon the results of the ANOVA and Newman-Keuls analyses, it was concluded that subjects were not able to discriminate between 32, 64, 128, and 256 shades of gray for static monochrome viewing. Because of this conclusion, it was decided to evaluate only one gray scale level from this group during the second evaluation of this study. Thirty-two shades of gray was selected for study, although any of the other three levels could have just as well been selected. Since subjects were consistently able to discriminate between 8 and 16 shades of gray and all the other levels, then these two gray scale levels were also included in this study's second evaluation. Therefore, Evaluation 2 studied the effects of 8, 16, and 32 shades of gray on operator perceptual discrimination performance.

EVALUATION 2

The objective of Evaluation 2 was to determine how color discrimination performance is affected by variant gray scale and worksite illumination conditions while viewing a monochrome monitor.

Subjects

Twelve volunteer subjects were selected to participate in this evaluation. Seven subjects were male and five were female. Subjects were randomly assigned to one of two different groups. These two different groups represented the different illumination conditions. All subjects who participated in Evaluation 2 had their vision tested at the JSC Clinic and it was determined that they all had either corrected or uncorrected 20/20 vision and none had evidence of color deficiencies.

Apparatus

Testing took place in the ROIL of the NASA Johnson Space Center. The video equipment used in the previous evaluation remained the same except that the DataCube was programmed to display gray scale levels of 8, 16, and 32. Two Lowel Omni 600 watt halogen lamps with a color temperature reading of 3200 degrees Kelvin were used so that the illumination levels could be varied. One of the lamps had spot lighting while the other had flood lighting.

Color chips from the *Munsell Book of Color* were used as the visual stimuli. Color chips selected were the 15 color chips determined by Frederick, Shields, and Kirkpatrick (1977) to be maximally discriminable for both color and direct viewing from a sample of 80 Munsell color chips evaluated. These chips with their respective hues, lightnesses and chromas are listed in Table 1.

The task performed in this evaluation was to determine from the image on the monitor whether two Munsell color chips were either the same or different from one another. The chips were placed on an off-white background that was remotely located away from the subject/monitor area. This background was deemed significant since it closely approximated the color of the payload bay, most thermal insulation, satellites, and structural members that are or will be used in space. Subjects viewed each possible two-chip combination of the 15 chips, with duplicate comparisons (two identical colors side by side) excluded from this study. Therefore, each subject viewed 105 paired-comparisons under each of the three gray scale conditions.

Variables

Two different independent variables were studied in this evaluation: monitor gray scale levels and worksite illumination conditions. The gray scale levels used were 8, 16, and 32 shades of gray. The two illumination levels, as measured at the location of the Munsell chips, consisted of a high illumination level of 16,021 lx (1489 fc) and a low level of 258 lx (24 fc). The high illumination level was approximately equivalent to the lighting conditions of the payload bay of the Shuttle with the sun shining at a high angle. The low illumination condition was approximately equivalent to the lighting conditions of the center of the payload bay of the Shuttle at night when flood lighting is used. Half of the subjects were randomly assigned to each of the two illumination groups. The design can thus be represented by a 2 x 3 two-factor repeated measures design with repeated measures on one factor -- the gray scale levels.

Procedure

The objectives of the evaluation and task instructions were briefly explained to the subjects. Subjects were instructed to respond to each paired-comparison by writing down on a data sheet whether or not the two chips viewed were either the same or different from one another. Each subject performed the 105 paired-comparisons for each of the three gray scale conditions for their respective level of illumination. Color combinations were presented at a rate such that subjects had three seconds in which to make a decision.

Partial counterbalancing was used. Three of the subjects in each group performed the paired-comparison task with the 8 gray scale condition viewed first, the 16 gray scale condition viewed second, and the 32 gray scale condition viewed third. The other three subjects in each group viewed 32 shades of gray first, 16 shades of gray second, and 8 shades of gray last.

Results and discussion

Each subject recorded responses for each of the 105 combinations while viewing all three gray scale levels for either the high or low illumination condition. In total, 315 data points were collected for each of the 12 subjects.

If a subject stated that any paired-comparisons were the same, then this was recorded as an error since none of the Munsell chips were compared to chips of an identical color. Figure 1 depicts the mean subject error-rates plotted across the three gray scale levels for both the high and low illumination conditions. This figure illustrates that the mean error-rates for the 16 gray scale condition was approximately the same (22.4 vs. 24.1) as the 32 gray scale while the mean error-rate for the 8 gray scale condition was substantially higher (37) -- for both illumination conditions.

These data were then statistically analyzed with an ANOVA. The analysis revealed that there was statistical significance ($p = 0.003$) due to the main effect of the different gray scales viewed. The results of the ANOVA did not reveal statistical significance due to the effect of the different illumination levels nor due to interaction effects.

Table 1. Munsell Color Chips used as stimuli.

HUE	COLOR	LIGHTNESS/ CHROMA	HUE	COLOR	LIGHTNESS/ CHROMA
2.5	red	4/14	7.5	green/yellow	6/12
3.75	red	4/14	7.5	green	5/10
8.75	red	max	7.5	green	4/10
6.25	yellow/red	max	7.5	blue/green	4/8
8.75	yellow/red	max	3.75	purple/blue	4/12
2.5	yellow	8/16	10	purple	5/12
2.5	green/yellow	7/12	10	purple	4/12
			5	red/purple	3/10

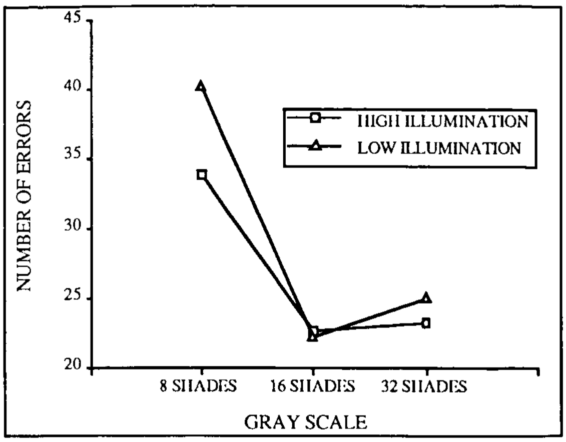

Figure 1. Paired-comparison errors plotted across gray scale and illumination levels

A Newman-Keuls pairwise comparison was then administered to these data. The analysis revealed that the subjects committed significantly more discrimination errors ($p < 0.05$) with the 8 gray scale condition than with either the 16 or 32 gray scale conditions. The Newman-Keuls did not show a significant difference between the error rates of the 16 and 32 gray scale viewing conditions. Since the error rate was significantly worse for viewing with 8 gray scales, it is recommended that for this particular task with monochrome viewing, at least 16 gray scales should be used. This is an interesting finding that may be an addition to the previously stated recommendations (Tannas, 1985) concerning the number of gray scale levels to use. It may well be that for static brightness-level discriminations, gray scale levels as low as 16 are sufficient.

After the study was conducted it was realized that if the most maximally discriminable colors from our test were identified, then these colors would also be discriminable under direct viewing and indirect color video viewing. This would be true because the fifteen color chips used in this study were found to be maximally discriminable for both direct view and indirect color video viewing by Frederick et al., 1975. This information might help establish guidelines concerning the selection of colors which have the least chance of being confused across all three possible viewing methods: direct viewing and both monochrome and color video viewing. An application for this information deals with methods of color coding materials which might be viewed directly by an EVA astronaut or remotely displayed on either monochrome or color video monitors.

When considering viewing under all three methods (direct, color video, and monochrome video), the driving factor is the monochrome condition. A literature review was conducted to determine how many brightness intensity levels are discriminable by an operator. The sources were found to be in disagreement on this matter. For example, NASA-STD-3000 states that no more than 3 brightness intensities be used, Engles and Granda (1975) stated that as many as 4 brightness intensities can be used with some risk of reduced legibility for the dimmer items, Grether and Baker (1972) recommend that no more than 4 levels be used, and Foley and Moray (1987) stated that between 3 and 5 absolute brightness discriminations can be made.

It was beyond the scope of this study to determine which recommendation is correct; therefore, the approach taken was to try to form all of the possible three, four, and five color combinations which had perfect discriminability between them. Since viewing with 8 gray scales was found to be unacceptable, then only the viewing under the 16 and 32 gray scale conditions was included in the analysis. The discrimination-error data across all 12 subjects and both lighting conditions was evaluated with the aid of computer analysis to form the combinations. This analysis yielded 93 three-color combinations, 9 four-color combinations, and 0 five-color combinations that yielded 100% discriminability within each specific color combination for these gray scale levels. The specific combinations may be found in Stuart, Bierschwale, and Smith (1989).

CONCLUSION

The results of this investigation determined that the perceptual discrimination of subjects performing a paired-comparison color-chip task under an 8 shades-of-gray viewing condition was significantly worse than their performance under 16 and 32 shades of gray. A statistically significant difference did not exist between 16 shades of gray and 32 shades of gray for this task. Even though the results in Evaluation 1 demonstrated that there is no perceptible

difference between 32 shades of gray and higher gray scale levels, the results obtained in this evaluation may not necessarily be generalizable to the performance of color discrimination tasks using monochrome viewing with 64, 128, and 256 shades of gray since the task was not performed with those levels.

Since the error rate was significantly worse for the 8 shades of gray condition, it is recommended that monitors used for remote monochrome viewing display at least 16 shades of gray if the tasks to be performed are perceptually similar to this task. If the remotely viewed task is more perceptually demanding than the task used in this evaluation, then the displayed gray scale value may need to be even higher. It is recommended that future investigations evaluate the effects of variant gray scale levels on more dynamic telerobotic tasks so that the results obtained will be more generalizable to telerobotic workstation design.

It was of interest to determine the number of three, four, and five-color combinations that had a 100% discriminability rate amongst themselves across both 16 and 32 gray scales and both lighting conditions. These are listed in color combination tables (Stuart et al., 1989) that can be used by systems designers who are faced with the question of how many colors and which specific colors to use for color-coding of tasks which will be performed either EVA or remotely performed and viewed through a monochrome and, or a color monitor. Even though there are other color combinations not evaluated in this study that would have probably produced perfect discriminability under these conditions, these data can eliminate the need to conduct an exhaustive evaluation.

This investigation has provided some insight into an important issue concerning the specific gray scale levels of the monitors to be used for monochrome viewing of a remote inspection task aboard *Freedom* and during later space-based activities. The major result of this study is that it has been determined that 8-gray-scale viewing is unsuitable for monochrome perceptual discrimination tasks. Another result of this study is that it has helped to establish color-coding guidelines concerning the colors which have the least chance of being confused with one another under variant gray scale and illumination conditions. These results have application to both telerobotic workstation monitors and coding of task hardware. The results also have relevance for the performance of remotely viewed tasks in the nuclear and undersea industries.

ACKNOWLEDGEMENTS

Support for this investigation was provided by the National Aeronautics and Space Administration through Contract NAS9-17900 to Lockheed Engineering and Sciences Company.

REFERENCES

Chandlee, G. O., Smith, R. L., and Wheelwright, C. D. (1988). Illumination requirements for operating a space remote manipulator. In W. Karwowski (Ed.), *Ergonomics of hybrid automated systems 1* (pp 241-248). Amsterdam: Elsevier.

Engel, S. E., and Granda, R. E. (1975). *Guidelines for man/display interfaces* (Technical Report TR 00.2720). Poughkeepsie, NY: IBM.

Foley, P. and Moray, N. (1987). Sensation, perception, and systems design. In G. Salvendy (Ed.), *Handbook of human factors* (pp 45-71). New York: John Wiley and Sons.

Frederick, P. N., Shields, N. L., and Kirkpatrick, Jr., M. (1977). *Earth orbital teleoperator visual system evaluation program* (Test Report Number 5). Huntsville, Alabama: Essex Corporation.

Grether, W. F., and Baker, C. A. (1972) Visual presentation of information. In H. P. Van Cott and R. G. Kinkade (Eds.), *Human engineering guide to equipment design* (pp 41-122). Washington, D. C.: American Institutes for Research.

Johnston, D. M. (1968). Target recognition on TV as a function of horizontal resolution and shades of gray. *Human Factors, 10*, 201-210.

Kingdom, F., and Moulden, B. (1986). Digitized images: What type of grey scale should one use? *Perception, 15*, 17-25.

NASA-STD-3000. (1987). *Man-system integration standards.* National Aeronautics and Space Administration.

O'Hara, J. M. (1986). *Telerobotic work system: Space-station truss-structure assembly using a two-arm dextrous manipulator* (Grumman Space Systems Report No. SA-TWS-86-R007). Bethpage, NY: Grumman Space Systems.

Shurtleff, D. A. (1980). *How to make displays legible.* La Mirada, CA: Human Interface Design.

Stuart, M. A., Bierschwale, J. M., and Smith, R. L. (1989). *Effects of variant gray scale levels and workplace illumination levels on operator performance.* (NASA Tech. Report JSC - 23483). Houston, Texas: NASA Lyndon B. Johnson Space Center.

Tannas, Jr., L. E. (1985). Flat-panel display design issues. In L. E. Tannas, Jr. (Ed.), *Flat-panel displays and CRTs* (pp 91-137). New York: Van Nostrand Reinhold.

Troy, E. B., Deutsch, E. S., and Rosenfeld, A. (1973). Gray-level manipulation experiments for texture analysis. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*, 91-98.

Woodson, W. E. (1981) *Human Factors Design Handbook.* New York: McGraw-Hill.